

# Through aggregation to deliberation? An experimental study on the effects of discussion summaries on users' willingness to comment and the deliberative quality of their contributions

Jana Leonie Peters, Marc Ziegele

Article - Version of Record

Suggested Citation:

Peters, J. L., & Ziegele, M. (2026). Through aggregation to deliberation? An experimental study on the effects of discussion summaries on users' willingness to comment and the deliberative quality of their contributions. *Computers in Human Behavior Reports*, 22, Article 101000.  
<https://doi.org/10.1016/j.chbr.2026.101000>

Wissen, wo das Wissen ist.



UNIVERSITÄTS- UND  
LANDESBIBLIOTHEK  
DÜSSELDORF

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20260414-103241-0>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>



# Through aggregation to deliberation? An experimental study on the effects of discussion summaries on users' willingness to comment and the deliberative quality of their contributions

Jana Leonie Peters<sup>\*</sup> , Marc Ziegele

Department of Social Sciences, Heinrich-Heine-University, Ulenbergstrasse 127, 40225, Duesseldorf, Germany

## ARTICLE INFO

### Keywords:

Online deliberation  
Online comments  
Information overload  
Knowledge

## ABSTRACT

Users' low willingness to participate in discussions in comment sections and the often-poor quality of their contributions have been identified as key challenges in online participation. To address these issues, previous research has proposed various strategies, including moderation. We argue that a less well-researched intervention, namely *aggregation* in the form of discussion summaries, reduces users' *information overload* and enhances their *objective knowledge* and *subjective knowledge*, which in turn are positively associated with their *willingness to participate* and the *deliberative quality of their comments*. Results from an online experiment ( $n = 643$ ) support most of our hypotheses, though objective knowledge does not directly impact willingness to comment. Differences between aggregation criteria were minimal, but fact-based aggregation was superior in improving objective knowledge compared to opinion- or argument-based approaches. These findings suggest that platform designers and moderators can utilize aggregation techniques to encourage participation and foster higher-quality online discourse.

## 1. Introduction

Online discussions in user comment sections on news websites and below publicly visible content on major social media platforms (e.g., Facebook, Instagram, YouTube, X) have the potential to enrich democratic discourse but often fall short of deliberative standards (Stroud et al., 2015). Only a minority of users actively engage in these discussions, and the quality of their contributions is frequently low, as they often lack argumentation and are marked by disrespectful behavior (Black et al., 2011; Esau et al., 2023; Papacharissi, 2004). A common explanation for the low quality is the unstructured nature of online discussions (Peacock et al., 2019). Scholars argue that the deliberative potential of the internet is not predetermined but co-shaped by the design and affordances of online platforms (Wright & Street, 2007). Consequently, several platform features have been explored for their potential to improve online discussions, including the availability of information (Gudowsky & Bechthold, 2013), the synchronicity of communication (Janssen & Kies, 2005; Strandberg & Berg, 2013), and moderation (Stroud et al., 2015; Wright & Street, 2007).

In addition to these approaches, aggregation is a promising design feature to foster participation and improve the quality of discussions in

comment sections. Aggregation involves restructuring discussions by summarizing (preliminary) results based on criteria such as arguments, opinions, or facts, and presenting these summaries prominently before the discussion (Peacock et al., 2019). Drawing on cognitive load theory (Sweller, 2010) and the Limited Capacity Model of Motivated Mediated Message Processing (Lang, 2000), aggregation can be understood as a mechanism that reduces informational complexity of discussions by distilling their essence. The effects of aggregation can be linked to established frameworks of information overload (Eppler & Mengis, 2004) and dual-process models of information processing (Chaiken, 1980; Petty & Cacioppo, 1986), which suggest that simplified yet structured information can enhance accessibility for users with varying levels of expertise. By abstracting from tone and style, aggregation can distill even uncivil contributions into their essential content, thereby reducing the visibility of inappropriate content and mitigating potential contagion effects (Kim et al., 2021).

Yet whether and how aggregation improves discussions in comment sections remains largely underexplored (except Peacock et al., 2019). Specifically, little is known about how different types of aggregation—such as summaries of arguments, opinions, or facts—affect users' willingness to participate and the quality of their contributions. The

<sup>\*</sup> Corresponding author.

E-mail addresses: [jana.leonie.peters@hhu.de](mailto:jana.leonie.peters@hhu.de) (J.L. Peters), [marc.ziegele@hhu.de](mailto:marc.ziegele@hhu.de) (M. Ziegele).

present study addresses this gap. Drawing on literature from media psychology and deliberation, we propose new mechanisms that link aggregation to reduced information overload and enhanced individual knowledge, two key factors influencing users' likelihood to comment and the quality of their contributions (Banducci et al., 2017; Kwon et al., 2015; Nematzadeh et al., 2019; Peacock et al., 2019; Ryffel & Wirth, 2020). Reducing users' cognitive load and increasing their knowledge through structured summaries could thus represent a pivotal intervention, particularly in high-volume comment sections where unstructured contributions can obscure valuable insights. By incorporating these psychological mechanisms into a comprehensive model of discussion aggregation effects, this study offers new insights to foster user engagement, promote high-quality discussions in comment sections, and provide implications for platform design.

## 2. Deliberation in comment sections

The concept of deliberative quality in online discussions is grounded in deliberative democracy, which emphasizes reasoned and respectful discourse as a foundation for democratic engagement (Friess et al., 2021; Habermas, 1996). Deliberation, according to Friess and Eilders (2015), refers to a type of communication in which participants exchange rational arguments, consider diverse perspectives, and aim for mutual understanding, reasoned disagreement, or consensus (*process perspective*). Ideally, a deliberative discussion space provides individuals with equal opportunities to express their viewpoints while maintaining a respectful and civil tone (Habermas, 1992, 1996). A high willingness of individuals to participate in the discourse is both a desirable *input* and a valuable *outcome* of deliberative discussions. It can be seen as a marker of inclusivity (Ziegele et al., 2020) as well as of user engagement and interactivity with the news (Ksiazek, 2018). Thus, in this study, users' willingness to participate in online discussions (input/outcome perspective) and the quality of their comments (process perspective) are used as indicators of deliberation.

Regarding these indicators, prior research has highlighted a tension between the inclusivity of participation and the quality of deliberation in online discussions. On the one hand, broad participation is desirable for democratic legitimacy and ensures that diverse perspectives are represented. On the other hand, encouraging higher-quality contributions often requires higher discursive standards, which may discourage participation, particularly among less confident or less knowledgeable users (Mutz, 2006; Stromer-Galley, 2007). This trade-off is central to debates on digital deliberation. Prior research has shown that challenges arise on both sides of this tension, with deficits in participation quantity as well as in the deliberative quality of contributions: only a small percentage of users actively participate (Stroud et al., 2016), and many discussions frequently lack substantive arguments and are marked by incivility (Anderson et al., 2014; Black et al., 2011; Coe et al., 2014; Ruiz et al., 2011). These findings are particularly concerning given that although few users write comments, far more users read them and form opinions based on the content (Ksiazek & Springer, 2018). User comments are viewed as a manifestation of social influence (Winter et al., 2015; Yun & Park, 2011), shaping perceptions of public opinion on specific topics (Lee & Jang, 2010), the perceived quality of news (Prochazka et al., 2018), and even influencing readers' commenting behavior (Kim et al., 2021; Springer et al., 2015).

Considering the broad readership and potential influence of discussions in comment sections, a growing body of research has examined the extent and quality of user participation. Influencing factors include the degree of deliberativeness and style of news articles (Ziegele et al., 2020), news factors (Weber, 2014), the nature of topics (Diakopoulos & Naaman, 2011; Tenenboim & Cohen, 2015), and the civility and quality of the comments posted by others (Cheng et al., 2017). Additionally, design features or affordances of social media platforms, such as moderation or asynchronous communication, affect both willingness to participate and the quality of contributions (Janssen & Kies, 2005;

Stroud et al., 2015). Among these factors, one promising yet underexplored design feature is aggregation (Faridani et al., 2010; Gao et al., 2017), which will be the focus of the next section.

## 3. Aggregation in comment sections

Although relatively few people write comments, the sheer volume of contributions—often hundreds or thousands per article—can become overwhelming (Faridani et al., 2010; Llewellyn et al., 2016). This poses challenges for maintaining deliberative quality, as not all comments contribute equally. Discussions in comment sections often feature a mix of constructive and unconstructive contributions due to diverse user backgrounds and levels of involvement (Hsu et al., 2009). Many comments are repetitive, superficial, or uncivil, increasing volume without advancing the discussion (Hargittai et al., 2012). An additional challenge is “reading overload” (Ma et al., 2012, p. 265), which arises when users confronted with vast discussions adopt selective reading strategies. Rather than engaging broadly, they often focus only on a few salient or extreme comments (Faridani et al., 2010), which can distort impressions of the debate. Research has shown that such biased reading allows polemic, uncivil, or superficial comments to dominate, gain visibility, and continue to attract attention, while other valuable insights are buried (Hsu et al., 2009). Reading overload also unfavorably shapes user attitudes and behaviors, often amplifying polarization, biased impressions of public opinion, or incivility (Anderson et al., 2014; Kim et al., 2021; Prochazka et al., 2018). Improving the conditions under which discussions in comment sections are read is therefore crucial not only for readers but also for the quality of subsequent user contributions.

To explain why aggregation may affect participation and deliberative quality in such contexts, this study draws on established models of information processing. Dual-process models distinguish between heuristic processing, which relies on simple cues and requires relatively little cognitive effort, and systematic processing, which involves deeper elaboration of arguments and factual information (Chaiken, 1980; Petty & Cacioppo, 1986). Aggregation can be understood as a design feature that directly shapes these processing conditions: by structuring and condensing information, aggregation may lower cognitive demands and facilitate heuristic processing, while at the same time enabling systematic processing by making central arguments, facts, or viewpoints more accessible. The dominating processing mode therefore depends not only on users' motivation, but also on how information is organized and presented within the discussion environment.

These assumptions are complemented by limited capacity approaches. This study refers specifically to the Limited Capacity Model of Motivated Mediated Message Processing (Lang, 2000), which extends general limited capacity models by explicitly incorporating motivational factors and by conceptualizing information processing as consisting of the subprocesses of encoding, storage, and retrieval. In unstructured comment sections, substantial cognitive resources are consumed by orientation and selection processes during encoding. Aggregation externalizes this organizational effort by pre-structuring the discussion, thereby freeing capacity for deeper encoding, storage, and retrieval. Thus, aggregation reallocates limited cognitive resources from navigation to understanding and deliberative engagement.

Aggregation synthesizes dispersed contributions into concise summaries of arguments, facts, or opinions (Faridani et al., 2010; Peacock et al., 2019). It can be seen as a design mechanism that may help to reconcile the trade-off between inclusivity and deliberative quality outlined in the previous section: by lowering cognitive entry barriers through reduced information overload, it can foster broader participation, while its structuring and knowledge-activating functions may elevate the quality of contributions. Unlike moderation, which enforces norms by removing or responding to problematic content (Jhaver et al., 2019; Matias, 2019; Seering et al., 2017), or threading, which merely organizes discussions hierarchically (Backstrom et al., 2013), aggregation reorganizes the informational environment. By condensing and

highlighting content, it helps users grasp the core of a debate with less time and effort, thereby mitigating overload and improving the signal-to-noise ratio of online deliberation systems (Towne & Herbsleb, 2012). Moreover, because summaries integrate relevant perspectives regardless of popularity, they ensure that diverse viewpoints are represented and that substantive content is not buried beneath noise. Prior research shows that argument-based summaries can stimulate both the quantity and the quality of deliberative engagement (Peacock et al., 2019), underscoring aggregation's dual role in enhancing accessibility for readers and supporting higher-quality contributions from writers.

Against this background, this study explores the psychological mechanisms behind the effects of aggregation. Building on the Limited Capacity Model of Motivated Mediated Message Processing and dual-process models of information processing, we argue that aggregation reduces information overload while at the same time enhancing users' subjective and objective knowledge. These mechanisms are expected to increase participation and improve comment quality.

Additionally, we analyze the specific effects of different aggregation criteria, such as summarizing arguments, opinions, or facts (Llewellyn et al., 2016; Peacock et al., 2019). Fig. 1 illustrates the conceptual model underlying our study. The labels assigned to the individual phases (e.g., information overload, objective knowledge, subjective knowledge) are consistently used to structure the formulation of hypotheses and research questions in the following sections. This approach is intended to improve clarity and orientation by directly linking each hypothesis and research question to its respective position in the model. Importantly, we formulate hypotheses where prior theory or empirical evidence allow for clear directional assumptions, while research questions are used where findings are inconsistent or insufficient.

3.1. Information overload, deliberation, and aggregation

The concept of information overload is closely tied to the limitations of human cognitive capacity. Working memory can process only a limited amount of information at a time (Miller, 1956; Sweller & Chandler, 1994). When information processing demands exceed these capacities, individuals experience information overload, which impairs their ability to process, store, and recall information effectively (Graf & Antoni, 2020). This phenomenon has both quantitative dimensions (volume of information) and qualitative dimensions (complexity, ambiguity, and relevance) (Eppler & Mengis, 2004).

In online discussions, these cognitive limitations have direct consequences for participation and deliberative quality. Research indicates that users experiencing overload are less likely to comment or engage constructively (Jones & Rafaeli, 1999; Peacock et al., 2019). During periods of low information load, users produce longer, more diverse messages, often incorporating questions and direct interactions

(Nematzadeh et al., 2019). Conversely, high-load conditions result in shorter, more stereotypical messages, reduced interaction, and increased reliance on emoticons (Jones et al., 2004; Nematzadeh et al., 2019). This creates a "cacophony" of overlapping voices that undermines meaningful communication (Nematzadeh et al., 2019, p. 10). Users experiencing overload may also disengage entirely or reduce their communication to superficial exchanges (Chen et al., 2009; Jones & Rafaeli, 1999).

Aggregation has the potential to alleviate the effects of information overload by summarizing and organizing complex discussions into more digestible formats (Chen et al., 2009; Faridani et al., 2010). By distilling key points and filtering irrelevant details, aggregation reduces cognitive burden and helps users focus on the most critical aspects of a discussion. While direct empirical evidence in the context of comment sections is limited, findings from e-commerce suggest that product summaries—such as aggregated reviews—reduce perceived overload (Hu & Krishen, 2019; Kwon et al., 2015; Park et al., 2006). We assume that these findings are applicable in comment sections.

Beyond reducing volume, aggregation also improves the quality of presented information. Low-quality information, characterized by ambiguity, irrelevance, or complexity, exacerbates overload (Hu & Krishen, 2019; Schneider, 1997). By contrast, high-quality, well-structured information reduces overload by enhancing users' processing fluency, even in large quantities (Sweller, 2010). Aggregation improves the "signal-to-noise ratio" (Hargittai et al., 2012, p. 162), making it easier for users to process and engage with the content.

In sum, we propose that aggregation increases users' willingness to participate and enhance the deliberative quality of their contributions through the reduction of information overload.

- H1.1. Discussion summaries in comment sections reduce users' information overload.
- H1.2. The less information overload users experience in comment sections, the higher their *willingness to comment*.
- H1.3. The less information overload users experience in comment sections, the higher the *deliberative quality of their comments*.
- H1.4. Discussion summaries in comment sections indirectly increase users' *willingness to comment* via reduced information overload.
- H1.5. Discussion summaries in comment sections indirectly increase *comment quality* via reduced information overload.

3.2. Knowledge, deliberation, and aggregation

Knowledge plays a central role in deliberation, influencing both users' willingness to participate and the quality of their contributions in

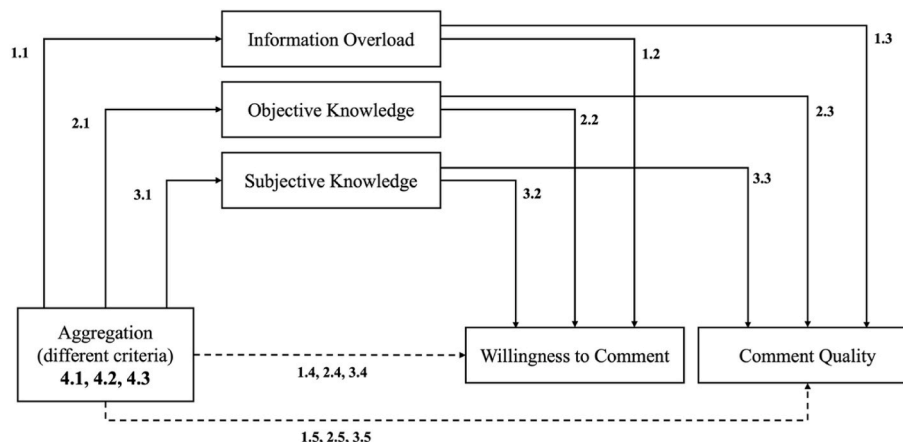


Fig. 1. Conceptual Model. The figure illustrates the labeling of the different paths that structure our theoretical framework.

online discussions. Knowledge can be understood along two dimensions: factual understanding (objective knowledge) and users' perceived grasp of a topic (subjective knowledge) (Flynn & Goldsmith, 1999; Schaefer, 2022). While objective knowledge involves concrete facts and the ability to integrate them meaningfully (Eveland & Hutchens, 2011; In der Au, 2017), subjective knowledge reflects users' self-assessment of their knowledge and confidence to engage (Nelson & Narens, 1990; Rios et al., 2017). Although these dimensions often overlap, discrepancies may occur, producing phenomena such as the "illusion of knowing," where perceived knowledge outpaces actual understanding (Glenberg et al., 1982). Subsequent work in metacognition research has explained this discrepancy by pointing to processing fluency: when information feels easy to process, individuals often believe they understand it better than they actually do (Alter & Oppenheimer, 2009; Bjork et al., 2013; Dunning et al., 2003; Koriat, 2012; Kruger & Dunning, 1999). Subjective knowledge is often linked to greater willingness to engage in discussions (Dalisy et al., 2012; Salmon & Neuwirth, 1990; Schaefer, 2020; Schaefer et al., 2022; Schneider et al., 2016). It also enhances users' self-confidence, a critical factor in deciding to contribute to discussions (Ran et al., 2016; Rios et al., 2017). Looking at objective knowledge, users with a solid factual knowledge face fewer difficulties in forming reasoned opinions and well-founded arguments (Shamir, 1997; Weber, 2014). However, while objective knowledge is commonly seen as a prerequisite for engagement (Legrand et al., 2019), its effect is often indirect and depends on factors such as topic familiarity or interest (Abdulrauf et al., 2017; Kim & Khang, 2014).

Knowledge also plays a critical role in shaping comment quality. Users with greater factual knowledge are more likely to provide thoughtful and insightful contributions to discussions (Heath & Douglas, 1990). Similarly, subjective knowledge can motivate users to aim for higher-quality contributions, although it does not guarantee adherence to deliberative norms (Ran et al., 2016). Discrepancies between subjective and objective knowledge may result in overconfidence, potentially impeding deliberative quality (Ryffel & Wirth, 2020), although this argument is yet to be empirically tested.

Again, we assume that aggregation can enhance both objective and subjective knowledge. By organizing and distilling key information, aggregation reduces cognitive load, improves information accessibility, and increases processing fluency (Banducci et al., 2017; Lang, 2000). Higher fluency facilitates the encoding, storage, and retrieval of information, supporting the acquisition of knowledge (Lang, 2000; Ryffel & Wirth, 2020). Additionally, the ease-of-processing heuristic suggests that fluently presented information is perceived as more memorable, boosting subjective knowledge (Kornell et al., 2011). The enhanced fluency of processing aggregated information, however, may also exacerbate the gap between objective and subjective knowledge, potentially reinforcing the "illusion of knowing" (Ryffel & Wirth, 2020). This dynamic underscores the importance of examining both dimensions of knowledge.

In sum, based on our review, we propose the following hypotheses and research questions regarding direct and indirect effects of aggregation on willingness to comment and comment quality via objective and subjective knowledge:

**H2.1.** Discussion summaries in comment sections increase users' objective knowledge.

**RQ2.2:** How is users' level of objective knowledge associated with their *willingness to comment*?

**H2.3.** The higher users' objective knowledge, the higher the *deliberative quality of their comments*.

**RQ2.4:** Is there an indirect positive effect of discussion summaries in comment sections on users' *willingness to comment* via objective knowledge?

**H2.5.** Discussion summaries in comment sections indirectly improve *comment quality* via increased objective knowledge.

**H3.1.** Discussion summaries in comment sections increase users' subjective knowledge.

**H3.2.** The higher users' subjective knowledge, the higher their *willingness to comment*.

**RQ3.3:** How is subjective knowledge associated with the *deliberative quality of comments*?

**H3.4.** Discussion summaries in comment sections indirectly improve *willingness to comment* via increased subjective knowledge.

**RQ3.5:** Is there an indirect positive effect of discussion summaries in comment sections on *comment quality* via subjective knowledge?

### 3.3. Aggregation criteria: facts, arguments, and opinions

Since aggregation in comment sections has rarely been studied or implemented, there is limited insight into the effects of different aggregation criteria. Previous work has shown that discussions can be aggregated based on different criteria, including arguments, facts, or opinions (Llewellyn et al., 2016; Peacock et al., 2019). Building on the information-processing perspective introduced above, these aggregation criteria can be understood as shaping distinct processing conditions by emphasizing different types of content and lowering cognitive demands.

Each criterion emphasizes different cognitive and communicative functions: facts provide direct access to verifiable information and may therefore facilitate objective knowledge acquisition by supporting accurate encoding and retrieval (Lang, 2000); arguments present competing perspectives, prompting cognitive elaboration and reasoned evaluation of trade-offs (Kuhn et al., 1995; Peacock et al., 2019; Petty & Cacioppo, 1986); and opinions convey subjective viewpoints and may foster identification and emotional resonance, potentially strengthening perceived understanding and engagement (Hubler, 2021).

These mechanisms have distinct implications for perceived information overload. Fact-based summaries may lower overload by offering structured access to key knowledge; argument-based summaries may increase cognitive complexity but simultaneously stimulate deliberation; and opinion-based summaries may simplify participation by signaling social cues, even if they are less informationally dense. In this sense, the effects of different aggregation criteria can again be understood in light of established models of information processing, such as the Limited Capacity Model of Motivated Mediated Message Processing (Lang, 2000) and dual-process approaches (Chaiken, 1980; Petty & Cacioppo, 1986).

These theoretical distinctions also translate into practical considerations for platform design. If the aim is to boost factual knowledge, fact-based summaries may be preferable; if the goal is to encourage deliberation, argument-based summaries could be more effective; and if the focus is on lowering participation barriers, opinion-based summaries might play a stronger role. However, as empirical research on these assumptions is sparse, we pose the following research questions:

**RQ4.1:** Do aggregated arguments, opinions, and facts differ in their effect on users' information overload?

**RQ4.2:** Do aggregated arguments, opinions, and facts differ in their effect on users' objective knowledge?

**RQ4.3:** Do aggregated arguments, opinions, and facts differ in their effect on users' subjective knowledge?

### 3.4. Summarizing assumed effects

Up to this point, we have presented several hypotheses and research questions regarding the relationships between information overload, objective and subjective knowledge, users' willingness to comment, and the deliberative quality of their comments. We also argued how discussion aggregation may affect these psychological mechanisms and thereby indirectly increase users' willingness to comment and the quality of their contributions. These hypotheses and research questions are integrated into the working model shown in Fig. 2. The model adopts the same phase labels as illustrated in Fig. 1 and specifies at which stages hypotheses (H) and research questions (RQ) were formulated.

## 4. Method

### 4.1. Research design and sample

This study was conducted as an online experiment, and we acknowledge our moral and professional responsibility to carry out this research ethically and responsibly. Participation was voluntary, and informed consent was obtained from all individual participants included in this study. The study fully complied with German legal standards and the European General Data Protection Regulation (GDPR/DSGVO), as well as established research ethics guidelines for empirical social research.

The ethical approval was waived by the Ethics Committee of the Arts and Humanities Faculty at Heinrich-Heine-Universität Düsseldorf, in accordance with German Research Foundation (DFG) guidelines.

In 2023, we recruited participants for a single-factor, between-subjects, web-based experiment via an external panel provider, resulting in a total sample of  $N = 670$  participants. Participants who did not pass a quality check—asking them to select a specific response option—were excluded. The final sample consisted of  $n = 643$  participants. All participants were residents of Germany and completed the study in German, situating the experiment within the context of the German political and media system.

The final sample included 48 percent female participants, with a mean age of 50 years ( $SD = 14.58$ ). Twenty-five percent had obtained a university entrance diploma, and 33 percent held an academic degree. Half of the participants completed the survey on a computer (laptop or desktop PC), and the other half on a mobile device (smartphone or tablet).

### 4.2. Procedure and stimulus material

After providing informed consent, participants completed socio-demographic questions and reported their opinion on cannabis legalization before being randomly assigned to one of four versions of an online discussion. The study was completed in a fixed sequence within a single online session.

We selected cannabis legalization as the topic because it had received sufficient media attention during the data collection period to ensure familiarity, while not being so extensively covered that participants already possessed deep knowledge. This approach intended to avoid floor or ceiling effects related to knowledge (Uttl, 2005; Wang et al., 2008). In addition, the topic was assumed to be sufficiently controversial to engage participants both cognitively and emotionally and to motivate them to comment but not so polarizing as to make them unreceptive to opposing arguments, facts, or opinions.

In three groups, the discussion was preceded by an aggregated summary. To achieve balanced group sizes for the main comparison of aggregation versus no aggregation, the control group without a summary was intentionally over-recruited to about three times its size. This resulted in roughly equal numbers of respondents with and without a summary, a strategy recommended for maximizing statistical power in two-group designs (Campbell et al., 1995). Within the aggregation condition, participants were evenly distributed across the three criteria—arguments, opinions, and facts—so that each subgroup was of comparable size while the discussion content remained identical across all groups.

All participants viewed a fictitious social media post on cannabis legalization. The post consisted of an image, a caption, a teaser text, a user discussion including 40 comments comprising a total of 1182 words of substantive comment text (approximately 30 words per comment on average), and, depending on the experimental condition, a discussion summary. This word count refers exclusively to the comment text and excludes aggregation, usernames, and other non-content elements. The comments were original user contributions posted on Facebook in response to an article by a nationwide news outlet. To ensure ethical compliance, only the text of these publicly accessible comments was used. They were manually transcribed rather than automatically crawled, and all personally identifying information (e.g., names, profile pictures, links) was removed. Where necessary, comments were slightly paraphrased or adjusted (e.g., adding sources or figures) to enable realistic aggregation while ensuring de-identification of the material and preventing identification of original authors by study participants.

Aggregation was operationalized by synthesizing recurring arguments, opinions, and facts from the original discussion into summaries

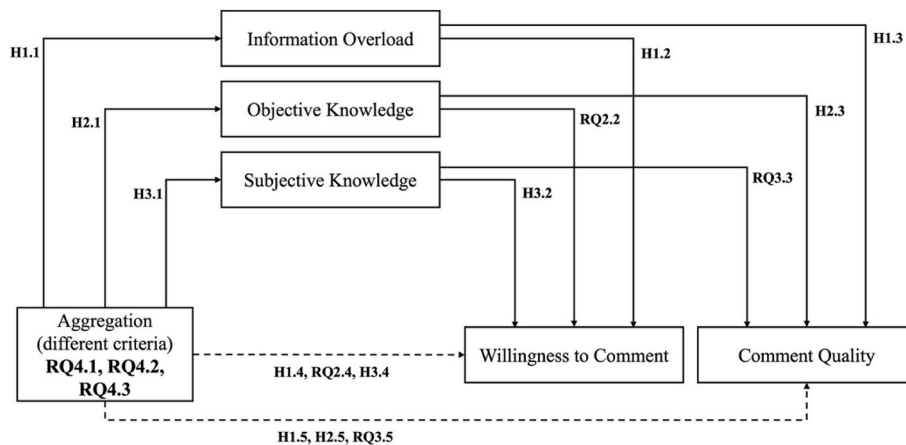


Fig. 2. Assumed Effects Model of Discussion Aggregation. The figure illustrates the assumed relationships between aggregation (based on different criteria), perceived information overload, objective and subjective knowledge, users' willingness to comment, and the deliberative quality of their comments. Solid arrows indicate direct effects, while dashed arrows represent indirect effects.

that were prominently positioned before the actual discussion (for an example, see Fig. 3).

In the “aggregated arguments condition”, the summary included six recurring arguments on cannabis legalization from the discussion (three pro and three con arguments). For instance, a pro argument stated that tobacco and alcohol are legal despite being more harmful than cannabis, while a con argument emphasized the potential health risks, such as psychosis.

In the “aggregated facts condition”, six recurring facts were summarized. Three facts supported legalization (e.g., 74,000 deaths annually in Germany from alcohol, 127,000 from tobacco, and none directly from cannabis), while three opposed it (e.g., experts claim cannabis triples the risk of psychosis and can harm young people’s brains). In our experiment, arguments and facts conveyed similar information in

different formats: arguments were explicitly labeled as pro or con according to their position, while facts included references to figures or experts. Opinions, in contrast, reflect subjective stances that cannot be directly tied to specific arguments or facts. In the “aggregated opinions condition”, seven recurring opinions ranging from pro to neutral to con were included, such as “Cannabis should only be legalized under strict regulations” or “Cannabis should not be legalized”.

During the initial exposure to the stimulus, participants were required to view the discussion for at least 45 s before they could proceed. After this minimum exposure time, they were free to spend as much time as they wished reviewing the stimulus before continuing. Participants could not return to the stimulus during subsequent questionnaire sections. On average, participants spent between 3.4 and 5.6 min viewing the stimulus, depending on condition (control group:  $M =$

7 hours ago

**The Traffic Light Coalition aims to fulfill its election promise and legalize cannabis despite opposition**



The planned legalization of cannabis in Germany is set to move forward. In August, the federal cabinet in Berlin approved the corresponding bill by Federal Health Minister Karl Lauterbach (SPD). However, the ministry does not expect the law to take effect until next year, as many questions remain unresolved.

317 comments

7 hours ago

**The Traffic Light Coalition aims to fulfill its election promise and legalize cannabis despite opposition**



The planned legalization of cannabis in Germany is set to move forward. In August, the federal cabinet in Berlin approved the corresponding bill by Federal Health Minister Karl Lauterbach (SPD). However, the ministry does not expect the law to take effect until next year, as many questions remain unresolved.

317 comments

**Summary | These are the key arguments from users:**

**Pro Legalization**

- Tobacco and alcohol are legal, even though they are more dangerous than cannabis
- Banning cannabis makes consumption more dangerous and doesn't prevent it
- Decriminalizing users would ease the burden on police and judiciary

**Against Legalization**

- Cannabis consumption can cause health issues like psychosis or brain damage
- Legalization would send the wrong message by trivializing consumption
- Cannabis use poses a danger to road traffic

---

**Patricia Bench**  
The first step would be to decriminalize smoking weed!!! If someone drives under the influence, of course, they should face the same penalties as before. But it can't be right that an 18-19 year old's life is made difficult just because they were caught with a bag of weed.

---

**Mehmet Demir**  
...I don't think that's very intoxicating LOL

---

**Alina M.**  
@Mehmet Demir I think it's very intoxicating haha

**These are the comments from users:**

---

**Patricia Bench**  
The first step would be to decriminalize smoking weed!!! If someone drives under the influence, of course, they should face the same penalties as before. But it can't be right that an 18-19 year old's life is made difficult just because they were caught with a bag of weed.

---

**Mehmet Demir**  
...I don't think that's very intoxicating LOL

---

**Alina M.**  
@Mehmet Demir I think it's very intoxicating haha

**Fig. 3.** Translated Treatment-Excerpts: “Aggregation: Arguments” and “No Aggregation”. The left panel shows the “aggregated arguments condition”, where a summary of three arguments in favor of and three against cannabis legalization precedes the user comments. The right panel displays the “no aggregation control condition”, in which users only see the same discussion thread without a preceding summary. Both treatments are based on the same mock news post and comment section. Importantly, the original stimuli contained 40 comments in all conditions.

4.5 min,  $SD = 5.9$ ; argument-based aggregation:  $M = 3.4$  min,  $SD = 2.8$ ; opinion-based aggregation:  $M = 5.6$  min,  $SD = 11.7$ ; fact-based aggregation:  $M = 4.7$  min,  $SD = 8.2$ ).

Prior to the main study, we conducted a pretest ( $N = 151$ , effective sample sizes ranged from  $n = 148$  to  $n = 149$  due to missing values) to ensure that the stimulus material was perceived as intended. The pretest was conducted with an independent sample, and none of the pretest participants took part in the main experiment. Participants in the pooled aggregation groups rated the stimuli as significantly better structured,  $F(1, 146) = 20.49$ ,  $p < .001$ ,  $\eta^2 = .12$  (4 items, 7-point scale,  $\alpha = .94$ ), more informative,  $F(1, 147) = 12.19$ ,  $p < .001$ ,  $\eta^2 = .08$  (4 items, 7-point scale,  $\alpha = .93$ ), and easier to comprehend,  $F(1, 147) = 12.86$ ,  $p < .001$ ,  $\eta^2 = .08$  (4 items, 7-point scale,  $\alpha = .93$ ), compared to the “no aggregation group”. As expected, no significant differences were observed in participants’ overall liking of the design,  $F(1, 146) = .02$ ,  $p = .882$ ,  $\eta^2 = .00$  (3 items,  $\alpha = .94$ ), or in how realistic they found the stimulus,  $F(1, 147) = 2.26$ ,  $p = .135$ ,  $\eta^2 = .02$  (3 items,  $\alpha = .82$ ). Details are reported in [Table A1.1](#) in the appendix. Additionally, [Table A1.2](#) in the appendix compares the three types of aggregated summaries (fact-, argument-, and opinion-based). The results demonstrate that they were rated similarly in terms of structure, informativeness, comprehensibility, design, and realism. This confirms that the manipulation differed only along the intended dimension, namely the type of information provided.

In the main study, after viewing the stimulus material, participants first completed an attention check assessing whether a discussion summary was present and, if applicable, which aggregation criterion (arguments, facts, or opinions) had been emphasized. Subsequently, participants responded to questions measuring perceived information overload as well as subjective and objective knowledge. The order of these measures was randomized.

After completing these measures, participants reported their willingness to comment. Immediately afterwards, they were prompted to write a comment. For this task, the full stimulus material (including the discussion and, where applicable, the aggregation summary) was displayed a second time, together with an open text field located below the comments. The input field did not visually replicate a specific platform interface. Importantly, participants were not required to write a comment and could skip this task and proceed with the survey. Finally, participants indicated how many of the displayed comments they had read, were given the opportunity to leave open-ended remarks, and completed the study with a debriefing.

The average completion time was 13.5 min ( $SD = 5.2$  min), and there were no upper time limits for any section apart from the minimum exposure time during initial stimulus presentation.

In sum, we aimed to ensure internal and external validity. Regarding external validity, the stimuli were designed to resemble real discussions, using mostly original comments. Results of the pretest indicate that participants generally perceived the content as realistic ( $M = 5.39$ ,  $SD = 1.46$ , 7-point scale) and authentic ( $M = 5.55$ ,  $SD = 1.45$ , 7-point scale), and stated that they would expect to encounter a similar display on a real website ( $M = 5.17$ ,  $SD = 1.61$ , 7-point scale).

### 4.3. Measures

Unless noted otherwise, all variables were measured on seven-point scales ranging from “1 = I do not agree at all” to “7 = I fully agree”.

#### 4.3.1. Information overload

Perceived information overload was measured with nine items ( $\alpha = .88$ ), for example “There was too much information”. The scale was based on previous research ([Hu & Krishen, 2019](#); [Kwon et al., 2015](#)) and supplemented with additional items reflecting the qualitative dimension of information overload (e.g., “The information was too complex”) ([Graf & Antoni, 2020](#); [Park et al., 2006](#)).

#### 4.3.2. Objective knowledge

Participants’ factual knowledge was assessed with five statements to be judged as true, false, or “I don’t know” (e.g., “Stretched cannabis often contains heroin”). All answers to these statements were available in each stimulus, but accessibility varied by condition: fact-based summaries provided some answers explicitly, argument-based summaries conveyed them more indirectly, and in the opinion-based and no-aggregation conditions, participants had to extract the information from the full set of comments. Each question had a 15-s response limit; unanswered items were automatically skipped and coded as missing values to prevent participants from looking up answers.

#### 4.3.3. Subjective knowledge

Subjective knowledge was measured with an eight-item scale ( $\alpha = .87$ ) adapted from prior research ([Schaefer et al., 2022](#)). Items captured both general knowledge (e.g., “I know quite a lot about the legalization of cannabis”) and stimulus-related knowledge (e.g., “The user discussion has made me feel well-informed about cannabis legalization”).

#### 4.3.4. Willingness to comment

Willingness to comment was measured with an 11-point scale. Participants rated how likely they would be to contribute a comment if they encountered the discussion online (0 = not at all likely, 10 = very likely). We deliberately used a wider response range to capture subtle differences in behavioral intention, given that willingness to join public online discussions is generally low. Prior research suggests that 11-point scales provide higher sensitivity and can enhance differentiation in self-report measures ([Lim, 2008](#)).

#### 4.3.5. Comment quality

Of the valid cases ( $n = 643$ ), 86 percent ( $n = 553$ ) of respondents posted a comment. After excluding spam (3%), 537 comments remained and were analyzed for deliberative quality. We applied quantitative content analysis along four dimensions ([Ziegele et al., 2020](#); [Stroud et al., 2015](#)): rationality (5 items, e.g., topic relevance, argument), reciprocity (2 items, e.g., reference to individual users, reference to all users), civility (9 items, e.g., explicit respect, absence of insults), expressive communication (5 items, e.g., positive emotions, storytelling). Intracoder reliability was calculated with ReCal ([Freelon, 2013](#)) using Krippendorff’s  $\alpha$  and was good to excellent across variables (.83–1.00), based on 65 coded comments. For intercoder reliability, a trained student assistant independently coded the same 65 comments. Krippendorff’s  $\alpha$  values were acceptable for most categories (.71–1.00), except for sarcasm, which reached only .66. Details are provided in [Tables A2 and A3](#) in the appendix.

The four dimensions were combined into a weighted additive index (similar to [Friess et al., 2021](#)) ranging from 0 (“not at all deliberative”) to 4 (“highly deliberative”). Each dimension contributed equally to the overall score. To achieve this balance, the variables within dimensions were weighted to account for differences in the number of items (e.g., 5 for rationality, 2 for reciprocity).

#### 4.3.6. Control variables

Participants’ age, sex, education, and political orientation were recorded at the beginning of the questionnaire. Opinions on cannabis legalization were assessed with a three-item scale (e.g., “The use of cannabis poses a high health risk”;  $\alpha = .85$ ). Participants also indicated how many of the displayed comments they had read, with response options ranging from “I read all of the comments” to “I did not read any comments.”

#### 4.3.7. Attention check

Participants were asked whether the discussion they had just read was preceded by a summary. If they answered “yes”, they were asked according to which criterion—arguments, opinions, or facts—the discussion was aggregated. In addition, participants indicated the extent to

which they agreed that a) arguments, b) opinions, or c) facts had been summarized under the article preview.

#### 4.4. Statistical analysis

We conducted mediation analyses using PROCESS (Hayes, 2021) in IBM SPSS 29. PROCESS applies the ordinary least squares method and calculates both standardized and unstandardized path coefficients for the total and direct effects, as well as unstandardized path coefficients for the indirect effects. All variables were entered into the analyses using their original scale metrics. No rescaling or standardization of the 7-point and 11-point measures was applied prior to analysis. To facilitate interpretation across differently scaled variables, standardized regression coefficients ( $\beta$ ) are reported alongside unstandardized estimates.

Bootstrapping with 5000 iterations and heteroscedasticity-consistent standard errors was applied (Davidson & MacKinnon, 1993). Indirect effects were considered significant if the confidence interval did not include zero.

Cases with missing values on the relevant variables were excluded listwise from the respective analyses. For objective knowledge, 32 participants had to be excluded, possibly due to time constraints in completing the respective items. For comment quality, 106 participants were excluded because they chose not to provide a comment or because their comment was classified as spam. Spam was defined as responses that contained no identifiable or interpretable content (e.g., single characters, nonsensical strings like xxxxaa232, test entries, or irrelevant filler text), as well as comments in which participants merely stated that they did not wish to write a comment.

No missing data occurred for the other constructs and variables.

To assess whether exclusions were systematically related to experimental conditions, we conducted chi-square tests aligned with the analytical strategy of the study. For the comparison between aggregation and no aggregation, exclusion due to missing objective knowledge did not differ between conditions,  $\chi^2(1) = .81, p = .369, V = .04$ , nor did exclusion from comment quality analyses,  $\chi^2(1) = 2.05, p = .152, V = .06$ . Within the aggregation conditions, exclusions were also not associated with aggregation type for objective knowledge,  $\chi^2(2) = 2.11, p = .349, V = .08$ , or comment quality,  $\chi^2(2) = 5.39, p = .068, V = .13$ . These results indicate that excluded cases were not systematically concentrated in specific conditions.

## 5. Results

### 5.1. Preliminary analyses

#### 5.1.1. Attention check

Most participants correctly identified whether they had been exposed to a discussion summary. A chi-square test showed a significant association between group membership and participants' reports of receiving a summary (see Table 1).

To examine whether participants recognized which criterion had been used to summarize the discussion, we conducted single-factor between-subjects analyses of variance, followed by post-hoc tests (see Table 2). Results indicate that participants accurately recognized opinion aggregation, but had difficulties distinguishing arguments and facts. This pattern likely reflects the conceptual overlap between factual and argumentative content in real-world discussions, which makes perfect separation difficult. Still, even where differences were not statistically significant, the strongest associations tended to occur in the theoretically expected direction, with the intended criterion being rated highest.

#### 5.1.2. Randomization checks

There were no significant differences between the aggregation groups and the no aggregation group in terms of their political left-right

**Table 1**  
Attention Check: Perception of Aggregation (No aggregation/Aggregation).

	No aggregation <i>n</i> = 332	Aggregation <i>n</i> = 311	Chi-square test			
	<i>n</i> (%)	<i>n</i> (%)	$\chi^2$	<i>df</i>	<i>p</i>	<i>V</i>
Perception of no aggregation	246 (74.1)	85 (27.3)	144.86	2	<.001	.48
Perception of aggregation	72 (21.7)	209 (67.2)				
No perception (I don't know)	14 (4.2)	17 (5.5)				

Note. Column percentages. *n* = absolute frequency, % = relative frequency,  $\chi^2$  = Chi-square, *df* = degrees of freedom, *p* = significance level, *V* = Cramer's V.

orientation,  $F(1, 641) = .03, p = .855, \eta^2 = .00$ , gender,  $\chi^2(1) = .36, p = .547, V = .02$ , or education level,  $\chi^2(3) = 3.40, p = .335, V = .07$ , as determined by ANOVAs or chi-square tests, depending on the level of measurement. A similar analysis for the individual aggregation criteria again revealed no significant differences for age,  $F(2, 308) = .06, p = .944, \eta^2 = .00$ , political left-right orientation,  $F(2, 308) = 1.76, p = .174, \eta^2 = .01$ , gender,  $\chi^2(2) = .59, p = .745, V = .04$ , or education level,  $\chi^2(6) = 11.91, p = .064, V = .07$ .

Details regarding the randomization checks are provided in the Appendix (Tables A4 to A9).

### 5.2. Hypotheses testing

#### 5.2.1. Information overload

Detailed results, including all relevant coefficients, are reported in Tables 3 and 4. As predicted, participants who experienced less information overload were more likely to comment (H1.2) and wrote higher-quality comments (H1.3). Regarding the effects of aggregation, participants who received discussion summaries reported significantly less information overload than those who did not, supporting H1.1. The mediation hypotheses proposing indirect effects of discussion summaries on willingness to comment (H1.4) and comment quality (H1.5) via information overload were also supported.

#### 5.2.2. Knowledge

Detailed results are reported in Tables 5–8. RQ2.2, which asked about the association between users' objective knowledge and their willingness to comment, showed no significant association between the two variables. Supporting H3.2, participants with higher subjective knowledge were more willing to comment. Additionally, participants with higher objective knowledge wrote comments of higher deliberative quality, supporting H2.3. Similarly, higher subjective knowledge was associated with a better deliberative quality in comments, addressing RQ3.3.

Regarding the effects of aggregation, hypotheses H2.1 and H3.1 were also supported: participants who received discussion summaries demonstrated higher objective knowledge and perceived themselves to be more knowledgeable than those who did not receive a summary.

The effect of discussion summaries on willingness to comment was fully transmitted via subjective knowledge: the indirect effect was significant, whereas the direct effect was no longer significant once subjective knowledge was included in the model, supporting H3.4. In addition, we found indirect effects of discussion summaries on comment quality via both objective knowledge (H2.5) and subjective knowledge (RQ3.5). By contrast, no indirect effect was observed on willingness to comment via objective knowledge (RQ2.4).

**Table 2**  
Attention check: Perception of aggregation criteria.

	Arguments <i>n</i> = 60	Opinions <i>n</i> = 80	Facts <i>n</i> = 69	ANOVA		
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>F</i> ( <i>df1</i> , <i>df2</i> )	<i>p</i>	$\eta^2$
Perception of arguments	<b>5.12 (1.83)</b>	4.43 (2.13)	4.45 (2.01)	2.46 (2, 206)	.088	.01
Perception of opinions <sup>1</sup>	4.78 (2.07) <sup>a</sup>	<b>5.79 (1.55)<sup>b</sup></b>	4.67 (1.98) <sup>a</sup>	9.27 (2, 125.54)	<.001	.07
Perception of facts <sup>1</sup>	4.92 (1.93) <sup>b</sup>	4.24 (2.14) <sup>a</sup>	<b>5.01 (1.74)<sup>b</sup></b>	3.30 (2, 134.09)	<.05	.03

Note. *M* = Mean, *SD* = Standard Deviation, *p* = significance level,  $\eta^2$  = Eta-Squared. Scale from 1 to 7. Groups without a common letter significantly differ from each other according to Games-Howell post-hoc tests (*p* < .05).

<sup>1</sup> Due to a significant Levene's test, Welch tests and Games-Howell post-hoc tests (*p* < .05) are reported.

**Table 3**  
Mediation analysis: Information overload – willingness to comment.

	<i>R</i> <sup>2</sup>	<i>b</i>	<i>SE</i> ( <i>HC3</i> )	$\beta$	<i>t</i>	<i>p</i>
<i>Outcome: Information overload</i>						
Constant		3.43	.07		49.75	<.001
Aggregation (path a)		-.21	.10	-.17	-2.12	.035
<i>Outcome: Willingness to comment</i>						
Constant	.029	3.34	.39		8.61	<.001
Aggregation (path c')		.48	.24	.16	2.03	.043
Information overload (path b)		-.34	.10	-.14	-3.53	<.001
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.07	.04	.003	.17

Note. *n* = 643. *b* = unstandardized regression coefficient, *SE* (*HC3*) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993),  $\beta$  = standardized regression coefficient, *p* = significance level; *BOOT SE* = bootstrap standard error, *CI* = confidence interval, *LB* = lower bound, *UB* = upper bound, 95% confidence level.

### 5.3. The effects of different aggregation criteria

We further investigated whether aggregation criteria differ in their effects on information overload, objective knowledge, and subjective knowledge. Overall, Table 9 indicates that aggregation criteria made little difference for information overload and subjective knowledge, but differed for objective knowledge. Specifically, objective knowledge was highest when the discussion was summarized based on central facts.

This finding aligns with the design rationale: fact-based summaries foregrounded several testable facts, whereas argument-based summaries referenced them more indirectly, and opinion-based summaries did not enhance accessibility relative to the control. The fact condition therefore provides a strong test of accessibility effects, while the other two conditions suggest that the potential benefits of aggregation may also result from improved processing rather than salience alone.

## 6. Discussion

This study investigated how and why aggregation in the form of discussion summaries influences user behavior in comment sections. Drawing on research in deliberation and media psychology, we developed a working model explaining how aggregation affects users' willingness to comment and the deliberative quality of their contributions via reductions in perceived information overload and increases in objective and subjective knowledge. Data from an online experiment

**Table 4**  
Mediation analysis: Information overload – comment quality.

	<i>R</i> <sup>2</sup>	<i>b</i>	<i>SE</i> ( <i>HC3</i> )	$\beta$	<i>t</i>	<i>p</i>
<i>Outcome: Information overload</i>						
Constant		3.40	.07		45.93	<.001
Aggregation (path a)		-.23	.11	-.18	-2.09	.038
<i>Outcome: Comment quality</i>						
Constant	.031	1.26	.03		37.76	<.001
Aggregation (path c')		.05	.02	.23	2.65	.008
Information overload (path b)		-.02	.01	-.12	-2.55	.011
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.01	.00	.0001	.01

Note. *n* = 537. *b* = unstandardized regression coefficient, *SE* (*HC3*) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993),  $\beta$  = standardized regression coefficient, *p* = significance level; *BOOT SE* = bootstrap standard error, *CI* = confidence interval, *LB* = lower bound, *UB* = upper bound, 95% confidence level.

supported most of our hypotheses. In the following, we discuss our findings in relation to theoretical and empirical work.

First, a reduction in perceived information overload was associated with both greater willingness to comment and higher comment quality. By decreasing the cognitive effort required to process large amounts of unstructured information, summaries enabled participants to engage more effectively in comment sections. Lower cognitive load freed up mental resources that could be directed toward formulating more thoughtful comments. This aligns with prior work showing that cognitive overload undermines engagement and contribution quality (e.g., Jones et al., 2004; Nematzadeh et al., 2019). From a practical perspective, our findings suggest that aggregation can increase both user activity and the quality of their contributions, potentially easing the burden on moderators to manage inappropriate content.

Second, objective knowledge was mainly linked to higher comment quality: participants with stronger factual understanding produced more thoughtful and well-reasoned contributions, corroborating earlier findings (Edwards et al., 2008). Consistent with prior research (Schaefer et al., 2022), however, objective knowledge did not directly predict willingness to comment. This suggests that factual knowledge is not a primary driver of participation, but when users do contribute, it enables them to provide more substantiated comments. Subjective knowledge, in contrast, was positively associated with both outcomes. Participants who perceived themselves as knowledgeable were more likely to engage, potentially driven by their confidence in their understanding of

**Table 5**  
Mediation analysis: Objective knowledge – willingness to comment.

	R <sup>2</sup>	b	SE (HC3)	β	t	p
<i>Outcome: Objective knowledge</i>						
Constant		3.21	.06		50.20	<.001
Aggregation (path a)		.35	.09	.31	3.89	<.001
<i>Outcome: Willingness to comment</i>						
Constant		1.76	.38		4.61	<.001
Aggregation (path c')		.61	.25	.20	2.46	.014
Objective knowledge (path b)		.12	.11	.04	1.06	.289
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.04	.04	-.03	.13

Note. n = 611. b = unstandardized regression coefficient, SE (HC3) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993), β = standardized regression coefficient, p = significance level; BOOT SE = bootstrap standard error, CI = confidence interval, LB = lower bound, UB = upper bound, 95% confidence level.

**Table 6**  
Mediation analysis: Objective knowledge – comment quality.

	R <sup>2</sup>	b	SE (HC3)	β	t	p
<i>Outcome: Objective knowledge</i>						
Constant		3.27	.07		47.73	<.001
Aggregation (path a)		.32	.10	.29	3.36	<.001
<i>Outcome: Comment quality</i>						
Constant		1.12	.03		37.83	<.001
Aggregation (path c')		.05	.02	.24	2.69	.007
Objective knowledge (path b)		.02	.01	.10	2.22	.018
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.01	.00	.001	.02

Note. n = 512. b = unstandardized regression coefficient, SE (HC3) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993), β = standardized regression coefficient, p = significance level; BOOT SE = bootstrap standard error, CI = confidence interval, LB = lower bound, UB = upper bound, 95% confidence level.

the topic (Ran et al., 2016; Rios et al., 2017). These findings imply that perceptions of knowledge often influence behavior more directly than actual knowledge. This resonates with metacognition research on “illusions of knowing,” which shows that fluency can increase perceived understanding even when factual knowledge remains unchanged (Bjork et al., 2013; Dunning et al., 2003; Kruger & Dunning, 1999). Aggregation may therefore promote participation by boosting users’ subjective confidence. Interestingly, in our study, this increase in subjective knowledge was not associated with lower, but with higher comment

**Table 7**  
Mediation analysis: Subjective knowledge – willingness to comment.

	R <sup>2</sup>	b	SE (HC3)	β	t	p
<i>Outcome: Subjective knowledge</i>						
Constant		3.79	.07		55.40	<.001
Aggregation (path a)		.30	.10	.24	3.08	.002
<i>Outcome: Willingness to comment</i>						
Constant		-1.34	.38		-3.54	<.001
Aggregation (path c')		.28	.22	.09	1.26	.208
Subjective knowledge (path b)		.92	.10	.38	9.10	<.001
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.27	.09	.10	.47

Note. n = 643. b = unstandardized regression coefficient, SE (HC3) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993), β = standardized regression coefficient, p = significance level; BOOT SE = bootstrap standard error, CI = confidence interval, LB = lower bound, UB = upper bound, 95% confidence level.

**Table 8**  
Mediation analysis: Subjective knowledge – comment quality.

	R <sup>2</sup>	b	SE (HC3)	β	t	p
<i>Outcome: Subjective knowledge</i>						
Constant		3.85	.07		51.75	<.001
Aggregation (path a)		.26	.11	.21	2.45	.015
<i>Outcome: Comment quality</i>						
Constant		1.06	.03		31.78	<.001
Aggregation (path c')		.05	.02	.22	2.52	.012
Subjective knowledge (path b)		.03	.01	.18	3.84	<.001
			<i>Effect</i>	<i>BOOT SE</i>	<i>LB 95% CI</i>	<i>UB 95% CI</i>
Indirect effect (a × b)			.01	.00	.002	.02

Note. n = 537. b = unstandardized regression coefficient, SE (HC3) = heteroscedasticity-consistent standard error (Davidson & MacKinnon, 1993), β = standardized regression coefficient, p = significance level; BOOT SE = bootstrap standard error, CI = confidence interval, LB = lower bound, UB = upper bound, 95% confidence level.

quality. This suggests that subjective knowledge generated through aggregation may differ from superficial confidence gained through minimal cues such as headline reading (Schaefer et al., 2022). While promising, this distinction needs to be tested in further contexts to assess whether aggregation reliably produces beneficial forms of subjective knowledge.

Our observation that aggregation reduces overload while enhancing knowledge can be attributed to reduced cognitive load and increased

**Table 9**  
Effect of aggregation criteria on the mediators.

	Arguments <i>n</i> = 100	Opinions <i>n</i> = 105	Facts <i>n</i> = 106	ANOVA		
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>F</i> ( <i>df</i> <sub>1</sub> , <i>df</i> <sub>2</sub> )	<i>p</i>	$\eta^2$
Information overload	3.18 (1.27)	3.26 (1.28)	3.20 (1.31)	.12 (2, 308)	.887	.00
Objective knowledge <sup>c</sup>	3.40 (.99) <sup>a</sup>	3.20 (1.07) <sup>a</sup>	4.06 (.94) <sup>b</sup>	20.09 (2, 295)	<.001	.12
Subjective knowledge	4.20 (1.17)	3.91 (1.25)	4.15 (1.15)	1.77 (2, 308)	.127	.01

Note. *M* = Mean, *SD* = Standard Deviation,  $\eta^2$  = Eta-squared, *p* = significance level. Scale from 1 to 7. Groups without a common letter significantly differ from each other (*p* < .001).

<sup>c</sup>Varying sample size: Arguments: *n* = 95, Opinions: *n* = 103, Facts: *n* = 100. Different scale: 0 to 5.

processing fluency. Summaries manage complexity by making content more accessible in terms of both volume and difficulty (Lang, 2000; Ryffel & Wirth, 2020; Sweller, 2010). Our pretest data support this interpretation: participants who received a summary rated the discussion as more structured, clearer, and easier to navigate. Improved organization likely contributed to both lower cognitive load and greater fluency (Ryffel & Wirth, 2020; Sweller, 2010). According to the Limited Capacity Model of Motivated Mediated Message Processing, this freed cognitive resources for storage and retrieval (Lang, 2000). Thus, summaries enhanced both subjective and objective knowledge (Banducci et al., 2017; Schaefer, 2022).

Interestingly, our findings challenge the assumption that reduced overload results from reading fewer comments (e.g., Kwon et al., 2015). A subsequent analysis shows no significant difference in the number of comments read between conditions,  $\chi^2(4) = 3.43, p = .489, V = .07$ . Instead, summaries appear to improve the efficiency of processing large volumes of information, helping users navigate “crowded” discussions without feeling overwhelmed, suggesting that reduced overload is driven by better cognitive management rather than reduced exposure to information.

The type of aggregation criterion did not significantly affect information overload or subjective knowledge. For platform design, this suggests that the choice of aggregating around facts, arguments, or opinions may not be decisive for these outcomes. By contrast, we observed an effect notable for objective knowledge: participants’ factual knowledge increased most when the discussion was summarized based on facts. This advantage likely stems from the more direct and detailed presentation of factual information, which facilitated encoding and retrieval processes. Argument-based summaries referenced factual content more indirectly, while opinion-based summaries did not enhance accessibility relative to the control condition. Importantly, as shown in Section 5.2, aggregation as a whole—regardless of criterion—significantly improved subjective knowledge and reduced information overload compared to the no-aggregation condition. This indicates that while fact-based summaries are particularly effective for factual knowledge acquisition, all three formats provide comparable benefits when it comes to reducing cognitive strain and fostering users’ perceived understanding.

Taken together, our findings contribute to theoretical debates on online deliberation by showing how aggregation in comment sections can both reduce cognitive strain and enhance deliberative confidence. At the same time, they highlight that even low-effort design interventions can have meaningful effects on user engagement and comment quality, pointing to a promising pathway for scalable implementation through human- and AI-based summarization.

### 6.1. Limitations

Despite the insights provided in this study, several limitations should be noted. First, the observed effects, although statistically significant, were small. This may reflect the complexity of factors influencing both the mediators and dependent variables. In addition, the discussion summaries represented only a minor part of the overall stimuli, which included a headline, image, teaser, and 40 user comments. Nevertheless,

these modest effects are meaningful given the low-effort nature of the intervention (brief exposure to a short summary) and the complexity of the targeted outcomes. Even small shifts in willingness or quality can scale to substantial differences in large comment sections. As one of the first experimental studies on aggregation in comment sections, our work thus provides a foundation for examining how interface-level interventions shape deliberative discourse.

Second, our manipulation of aggregation versus no aggregation did not fully isolate the structural effect of summarization from the effect of providing additional information. Participants in the aggregation conditions received an extra text element (the summary), while the control group did not. This mirrors how aggregation is typically implemented on platforms (e.g., Peacock et al., 2019), but it also means that some of the observed effects—particularly for knowledge—may reflect information volume rather than structure alone. Future studies should therefore include control conditions with unstructured or randomly selected content to disentangle these effects. A related issue concerns objective knowledge: although all items appeared in the full set of comments, their salience varied across aggregation types. Fact-based summaries explicitly highlighted several items, whereas argument- and opinion-based summaries did not increase accessibility to the same extent. This may constrain comparability across conditions, even though it reflects the practical functioning of aggregation, which is designed to surface key points.

Third, the applicability of our results may depend on the issue under discussion. Cannabis legalization is a topic with relatively clear pro and con positions, which may have made it easier for participants to process information and articulate their views. Other issues may be more complex or less reducible to dichotomous positions, which could change how aggregation operates and alter its effects.

Fourth, some participants had difficulty distinguishing between the aggregation criteria (facts, arguments, opinions). This may reflect the conceptual overlap of these categories in real-world discussions and the limited salience of distinctions in our stimuli. At the same time, the between-condition effects consistently aligned with theoretical expectations, and participants rated the intended criterion highest in most cases. This suggests that, although the distinctions may not have been as salient as intended, participants likely perceived contrasts between the formats that, due to the small group sizes, may not have reached sufficient statistical power to achieve significance. Future studies may use more distinct or stylized summary formats to sharpen these contrasts.

Fifth, the experimental design carries limitations for external validity. Participants were prompted to write a comment after indicating their willingness, but they could also skip this task. While this encouraged engagement, it may not fully capture natural commenting behavior. Moreover, the absence of direct interaction with other users may have affected comment quality. Prompting may also have induced some participants to comment who would not have done so in a real setting. In practice, online discussions are shaped by self-selection, which our design cannot fully replicate. In addition, the study was conducted within the German political and media context. Norms of political participation, commenting behavior, and deliberative expectations may differ across countries and media systems, which may limit the generalizability of the findings beyond this national context.

Sixth, potential biases in measurement should be acknowledged. Subjective knowledge was assessed with self-report items, which may be prone to social desirability or response inflation. Commenting behavior and perceptions of overload may also have been influenced by task framing or the awareness of being studied (i.e., Hawthorne effects). A further limitation is that willingness to comment was measured on an 11-point scale, whereas most other constructs used 7-point scales. Although this choice was intentional to increase sensitivity, it may limit comparability across measures.

Ultimately, the index of deliberative comment quality weighted the four dimensions—rationality, reciprocity, civility, and expressive communication—equally, even though some dimensions contained more variables than others. This approach may have influenced the overall score and potentially underestimated effects on individual dimensions.

## 6.2. Implications & conclusion

This study highlights the potential of aggregation, in the form of discussion summaries, to improve online deliberation in comment sections by increasing users' willingness to comment and enhancing the quality of their contributions. By reducing information overload, aggregation makes complex discussions more accessible and reduces cognitive strain. Our findings indicate that summaries not only help users process information more effectively but also increase both objective and subjective knowledge, leading to more thoughtful and deliberative comments.

Theoretically, our research advances understanding of how online discussion design can mitigate information overload and foster meaningful user participation. By structuring contributions and highlighting key points, platforms can improve the deliberative quality of comments. Distinguishing between subjective and objective knowledge provides a more nuanced account of user engagement, underscoring the pivotal role of perceived knowledge in motivating participation and shaping comment quality.

For platform designers and moderators, aggregation techniques offer a way to foster more inclusive and higher-quality discourse. Summarizing key facts, arguments, or opinions, creates a more manageable environment, encouraging participation and raising the overall quality of public discussions. Since aggregation criteria did not significantly affect perceived overload, platforms can select whichever format best suits the context without compromising user experience.

In this study, summaries were created by human coders, which ensured deliberative relevance and fairness. Yet human curation is difficult to scale in high-volume comment sections. Advances in natural language processing and large language models now make automated summarization pipelines feasible (Bhaskar et al., 2023). While such systems face challenges of accuracy, neutrality, and representativeness (Cao et al., 2022; Li et al., 2025), their viability increases when combined with lightweight fact checks and selective human review.

Beyond questions of scalability, the aggregation of arguments, facts, or opinions in political discussions also raises concerns about perceived bias and trustworthiness, particularly when such aggregation is implemented at scale using AI-based systems. Even when summaries are designed to be neutral and comprehensive, users may question how content is selected, framed, or prioritized, which can affect trust in both the aggregation mechanism and the hosting platform. These concerns are especially salient in politically sensitive contexts, where perceptions of imbalance or selective emphasis may undermine deliberative legitimacy. Transparent aggregation criteria and clear communication about how summaries are generated therefore appear crucial for maintaining

user trust alongside the deliberative benefits identified in this study.

Our results provide a foundation for further exploration of aggregation in online deliberation. Future work should test whether effects generalize across topics, platforms, and audiences, particularly in debates that are more complex or less dichotomous than cannabis legalization. Longitudinal designs could assess whether improvements in knowledge, overload, and comment quality persist over time, while experiments in interactive settings (e.g., live chats, threaded forums) could capture more naturalistic dynamics. Comparative studies of human-versus AI-generated summaries, replications on issues such as climate policy or migration, and cross-topic robustness checks would help clarify topic sensitivity and boundary conditions.

In addition, future studies could move beyond composite measures of deliberative quality to examine subdimensions such as rationality, reciprocity, civility, and expressive communication individually. While we relied on an index for reasons of consistency and comparability, disaggregated analyses may reveal whether aggregation selectively fosters certain facets of deliberation, thereby offering a more differentiated understanding of its effects.

Finally, studies should examine how aggregation effects can be amplified, for example by combining summaries with interactive features (e.g., voting, highlighting), adapting them dynamically to user needs (e.g., layered or expandable formats), timing exposure strategically to avoid bias, or complementing text with visual aids such as argument maps (Roush & Balaji, 2020; Roush et al., 2024). Pursuing these avenues will clarify the boundaries of aggregation's effectiveness and inform the design of scalable deliberative systems.

Taken together, our findings show that aggregation can simultaneously reduce cognitive strain and strengthen deliberative confidence. They highlight how even low-effort design interventions can meaningfully shape user engagement and comment quality, offering a promising pathway for scalable implementation through AI-supported summarization.

## CRedit authorship contribution statement

**Jana Leonie Peters:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Conceptualization. **Marc Ziegele:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author(s) used *ChatGPT (OpenAI)* in order to improve the language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

## Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chbr.2026.101000>.

Appendix

**Table A1.1**  
Treatment Check: Perception of the Stimuli I

	No aggregation <i>n</i> = 43	Aggregation <i>n</i> = 105/106	ANOVA		
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>F</i> ( <i>df</i> <sub>1</sub> , <i>df</i> <sub>2</sub> )	<i>p</i>	$\eta^2$
<b>Structure (<math>\alpha = .94</math>)</b>	3.18 (1.63)	4.57 (1.72)	20.49 (1, 146)	<.001	.12
Well structured	3.12 (1.76)	4.44 (1.87)	15.89 (1, 147)	<.001	.10
Clear	3.33 (1.89)	4.50 (1.95)	11.28 (1, 147)	<.001	.07
Illustrative	3.42 (1.67)	4.87 (1.76)	21.20 (1, 146)	<.001	.13
Easy to find relevant information	2.86 (1.89)	4.41 (1.96)	19.43 (1, 147)	<.001	.12
<b>Informativeness (<math>\alpha = .93</math>)</b>	3.17 (1.74)	4.20 (1.59)	12.19 (1, 147)	<.001	.08
Informative	3.16 (1.83)	4.58 (1.71)	20.32 (1, 147)	<.001	.12
Provided new information	3.07 (1.93)	3.72 (1.78)	3.86 (1, 147)	.051	.03
Helpful to expand knowledge	2.88 (1.87)	4.01 (1.87)	11.10 (1, 147)	<.01	.07
Conveyed key aspects of the topic	3.56 (1.89)	4.49 (1.72)	8.48 (1, 147)	<.01	.06
<b>Comprehensibility (<math>\alpha = .93</math>)</b>	4.82 (1.52)	5.63 (1.13)	12.86 (1, 147)	<.01	.08
Easy to understand <sup>a</sup>	5.19 (1.52)	5.95 (1.09)	9.10 (1, 60.41)	<.01	.08
Clearly formulated	4.21 (1.86)	5.17 (1.53)	10.64 (1, 147)	<.01	.07
Easy to follow <sup>a</sup>	5.00 (1.62)	5.60 (1.26)	4.81 (1, 63.48)	<.05	.04
Understandable <sup>a</sup>	4.88 (1.55)	5.80 (1.17)	12.28 (1, 62.59)	<.001	.10
<b>Design (<math>\alpha = .94</math>)</b>	4.32 (1.64)	4.36 (1.64)	.02 (1, 146)	.882	.00
Appealing webpage design	4.09 (1.78)	4.30 (1.78)	.40 (1, 146)	.531	.00
Liking of the webpage	4.37 (1.76)	4.12 (1.80)	.59 (1, 146)	.444	.00
Consistent webpage design	4.49 (1.61)	4.67 (1.65)	.36 (1, 146)	.548	.00
<b>Realism (<math>\alpha = .82</math>)</b>	5.62 (1.26)	5.27 (1.30)	2.26 (1, 147)	.135	.02
Realistic presentation of information	5.44 (1.42)	5.37 (1.48)	.08 (1, 147)	.780	.00
Authentic content	5.72 (1.39)	5.48 (1.48)	.83 (1, 147)	.363	.01
Expectation of similar presentation on a real news site	5.70 (1.49)	4.96 (1.62)	6.59 (1, 147)	<.05	.04

Note. *M* = Mean, *SD* = Standard Deviation,  $\eta^2$  = Eta-squared, *p* = significance level. Scale from 1 to 7. Pretest-Data.

<sup>a</sup> Due to a significant Levene's test, Welch tests are reported.

**Table A1.2**  
Treatment Check: Perception of the Stimuli II

	Argument-based <i>n</i> = 34	Opinion-based <i>n</i> = 38	Fact-based <i>n</i> = 34	ANOVA		
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>F</i> ( <i>df</i> <sub>1</sub> , <i>df</i> <sub>2</sub> )	<i>p</i>	$\eta^2$
<b>Structure<sup>a</sup> (<math>\alpha = .94</math>)</b>	4.61 (1.73)	4.67 (1.48)	4.41 (1.98)	.20 (2, 65.32)	.823	.00
Well structured	4.44 (1.83)	4.71 (1.66)	4.15 (2.13)	.81 (2, 103)	.448	.02
Clear <sup>a</sup>	4.47 (2.11)	4.71 (1.59)	4.29 (2.18)	.45 (2, 103)	.642	.01
Illustrative	4.91 (1.77)	4.95 (1.61)	4.74 (1.96)	.14 (2, 102)	.869	.00
Easy to find relevant information	4.44 (1.98)	4.32 (1.74)	4.47 (2.22)	.06 (2, 103)	.939	.00
<b>Informativeness (<math>\alpha = .93</math>)</b>	4.35 (1.53)	3.92 (1.51)	4.36 (1.75)	.91 (2, 103)	.405	.02
Informative	4.68 (1.55)	4.34 (1.70)	4.76 (1.89)	.61 (2, 103)	.543	.01
Provided new information	3.97 (1.80)	3.32 (1.74)	3.91 (1.76)	1.54 (2, 103)	.220	.03
Helpful to expand knowledge	4.18 (1.83)	3.71 (1.86)	4.18 (1.93)	.75 (2, 103)	.473	.01
Conveyed key aspects of the topic	4.59 (1.56)	4.32 (1.54)	4.59 (2.06)	.30 (2, 103)	.740	.01
<b>Comprehensibility (<math>\alpha = .93</math>)</b>	5.58 (1.27)	5.55 (1.04)	5.78 (1.09)	.43 (2, 103)	.652	.01
Easy to understand	5.82 (1.31)	5.89 (1.06)	6.15 (.86)	.83 (2, 103)	.439	.02
Clearly formulated	5.12 (1.51)	5.21 (1.40)	5.18 (1.71)	.03 (2, 103)	.968	.00
Easy to follow	5.53 (1.46)	5.45 (1.10)	5.85 (1.18)	1.03 (2, 103)	.362	.02
Understandable	5.85 (1.31)	5.63 (1.15)	5.94 (1.07)	.67 (2, 103)	.516	.01
<b>Design (<math>\alpha = .94</math>)</b>	4.47 (1.62)	4.44 (1.59)	4.16 (1.75)	.36 (2, 102)	.701	.01
Appealing webpage design	4.44 (1.74)	4.32 (1.77)	4.12 (1.85)	.27 (2, 102)	.762	.00
Liking of the webpage	4.32 (1.74)	4.21 (1.66)	3.82 (2.01)	.73 (2, 102)	.485	.01
Consistent webpage design	4.65 (1.57)	4.79 (1.71)	4.55 (1.68)	.20 (2, 102)	.823	.00
<b>Realism (<math>\alpha = .82</math>)</b>	5.14 (1.57)	5.52 (1.11)	5.13 (1.20)	1.07 (2, 103)	.345	.02
Realistic presentation of information <sup>a</sup>	5.09 (1.64)	5.79 (1.19)	5.18 (1.53)	2.89 (2, 65.35)	.063	.05
Authentic content	5.35 (1.63)	5.68 (1.47)	5.38 (1.35)	.55 (2, 103)	.578	.01
Expectation of similar presentation on a real news site	4.97 (1.82)	5.08 (1.40)	4.82 (1.68)	.22 (2, 103)	.803	.00

Note. *M* = Mean, *SD* = Standard Deviation,  $\eta^2$  = Eta-squared, *p* = significance level. Scale from 1 to 7. Pretest-Data.

<sup>a</sup> Due to a significant Levene's test, Welch tests are reported.

**Table A2**  
Content Analysis: Operationalization

Dimension/Variable	Category	Definition	Based on
<b>Rationality</b>			
Topic relevance	[0] absent [1] present	Indicates whether a comment refers to the topic of the article preview (cannabis legalization).	Esau et al., 2023; Freelon, 2010; Ruiz, 2011
Argument	[0] absent [1] present	Indicates whether a comment contains at least one argument.	Esau, 2022; Esau et al., 2017, 2023; Kies, 2010; Stromer-Galley, 2007
Additional knowledge	[0] absent [1] present	Indicates whether a comment contains additional knowledge.	Ziegele et al., 2020
Genuine question	[0] absent [1] present	Indicates whether a comment contains genuine questions.	Esau, 2022; Esau et al., 2023; Stromer-Galley, 2007
Constructiveness	[0] absent [1] present	Indicates whether a comment constructively influences the debate.	Esau et al., 2017, 2023; Monnoyer-Smith & Wojcik, 2012
<b>Reciprocity</b>			
Reference to individual users	[0] absent [1] present	Indicates whether a comment refers to one or more users.	AUTHOR CITATION; Esau et al., 2017; Rowe, 2015
Reference to all users	[0] absent [1] present	Indicates whether a comment refers to all users.	AUTHOR CITATION; Esau et al., 2017; Rowe, 2015
<b>Civility</b>			
Explicit respect	[0] absent [1] present	Indicates whether a comment contains explicit expressions of respect, appreciation, or thanks.	AUTHOR CITATION; Steenbergen et al., 2003
Absence of insults	[0] absent [1] present	Indicates whether a comment is free of insults.	Esau et al., 2017
Absence of shouting	[0] absent [1] present	Indicates whether a comment is free of shouting.	Friess et al., 2021
Absence of vulgarity	[0] absent [1] present	Indicates whether a comment is free of vulgarity.	Friess et al., 2021
Absence of sarcasm/cynicism	[0] absent [1] present	Indicates whether a comment is free of sarcasm/cynicism.	Friess et al., 2021
Absence of lies/false factual claims	[0] absent [1] present	Indicates whether a comment is free of lies or false factual claims.	Friess et al., 2021
Absence of threats	[0] absent [1] present	Indicates whether a comment is free of threats.	Friess et al., 2021
Absence of advocacy for violence/negative consequences	[0] absent [1] present	Indicates whether a comment is free of advocacy for violence or negative consequences.	Friess et al., 2021
Absence of group related hostility	[0] absent [1] present	Indicates whether a comment is free of group-related hostility.	Friess et al., 2021
<b>Expressive Communication</b>			
Positive emotions	[0] absent [1] present	Indicates whether a comment contains positive emotions.	Esau, 2022; Esau et al., 2023
Negative emotions	[0] absent [1] present	Indicates whether a comment contains negative emotions.	Esau, 2022; Esau et al., 2023
Storytelling	[0] absent [1] present	Indicates whether a comment contains stories or personal experiences.	Black, 2008; Esau, 2022; Esau et al., 2023; Polletta & Lee, 2006
Innocuous humor	[0] absent [1] present	Indicates whether a comment contains jokes, humorous content, or light-heartedness.	Esau, 2022; Esau et al., 2023
Empathy	[0] absent [1] present	Indicates whether a comment acknowledges or empathizes with perspectives or feelings.	Esau et al., 2023; Kies, 2010

**Table A3**  
Content Analysis: Reliability

Dimension/Variable	Intracoder reliability (Krippendorff's $\alpha$ )	Intercoder reliability (Krippendorff's $\alpha$ )
<b>Rationality</b>		
Topic relevance	1	.81
Argument	.85	.77
Additional knowledge	.85	.82
Genuine question	.95	.89
Constructiveness	.83	.71
<b>Reciprocity</b>		
Reference to individual users	1	.93
Reference to all users	1	.89
<b>Civility</b>		
Explicit respect	1	.93
Absence of insults	.95	.79
Absence of shouting	1	.85
Absence of vulgarity	.95	.90
Absence of sarcasm/cynicism	1	.66
Absence of lies/false factual claims	1	.92
Absence of threats	1	1
Absence of advocacy for violence/negative consequences	1	.94
Absence of group related hostility	1	.87
<b>Expressive Communication</b>		

(continued on next page)

**Table A3 (continued)**

Dimension/Variable	Intracoder reliability (Krippendorff's $\alpha$ )	Intercoder reliability (Krippendorff's $\alpha$ )
Positive emotions	1	.84
Negative emotions	.94	.94
Storytelling	.89	.95
Innocuous humor	1	.91
Empathy	1	.95

**Table A4**

Randomization Check: No aggregation/Aggregation

	No Aggregation $n = 332$	Aggregation $n = 311$	ANOVA		
	$M (SD)$	$M (SD)$	$F (df1, df2)$	$p$	$\eta^2$
Left-right	4.78 (1.83)	4.76 (1.98)	.03 (1, 641)	.855	.00
Age <sup>a</sup>	50.12 (14.42)	50.17 (14.77)	.00 (1, 641)	.965	.00

Note.  $M$  = Mean,  $SD$  = Standard Deviation,  $\eta^2$  = Eta-squared,  $p$  = significance level. Scale from 1 to 7.

<sup>a</sup> Different value range: Open-ended question, Min = 18, Max = 74.

**Table A5**

Randomization Check: Aggregation Criteria

	Arguments $n = 100$	Opinions $n = 105$	Facts $n = 106$	ANOVA		
	$M (SD)$	$M (SD)$	$M (SD)$	$F (df1, df2)$	$p$	$\eta^2$
Left-right	4.59 (1.93)	4.62 (1.90)	5.05 (2.09)	1.76 (2, 308)	.174	.01
Age <sup>a,b</sup>	50.05 (14.66)	50.55 (13.64)	49.91 (16.04)	.06 (2, 204.31)	.944	.00

Note.  $M$  = Mean,  $SD$  = Standard Deviation,  $\eta^2$  = Eta-squared,  $p$  = significance level. Scale from 1 to 7.

<sup>a</sup> Different value range: Open-ended question, Min = 18, Max = 74.

<sup>b</sup> Due to a significant Levene's test, the Welch test is reported.

**Table A6**

Randomization Check: Gender (No aggregation/Aggregation)

	No aggregation $n = 332$	Aggregation $n = 311$	Chi-square test			
	$n (%)$	$n (%)$	$\chi^2$	$df$	$p$	$V$
Gender			.36	1	.547	.02
Female	157 (47.3)	152 (49.7)				
Male	175 (52.7)	154 (50.3)				

Note. Column percentages.  $n$  = absolute frequency, % = relative frequency,  $\chi^2$  = Chi-square,  $df$  = degrees of freedom,  $p$  = significance level,  $V$  = Cramer's V. "Other" excluded from the analysis.

**Table A7**

Randomization Check: Education (No Aggregation/Aggregation)

	No aggregation $n = 329$	Aggregation $n = 311$	$\chi^2$	$df$	$p$	$V$
	$n (%)$	$n (%)$				
Education			3.40	3	.335	.07
No degree	0 (.0)	0 (.0)				
Lower secondary school degree	38 (11.6)	25 (8.0)				
Intermediate secondary school degree	102 (31.0)	106 (34.1)				
(Technical) university entrance qualification	86 (26.1)	73 (23.5)				
University degree	103 (31.3)	107 (34.4)				

Note. Column percentages.  $n$  = absolute frequency, % = relative frequency,  $\chi^2$  = Chi-square,  $df$  = degrees of freedom,  $p$  = significance level,  $V$  = Cramer's V. "Still in school" excluded from the analysis.

**Table A8**

Randomization Check: Gender (Aggregation Criteria)

	Arguments $n = 98$	Opinions $n = 105$	Facts $n = 103$	$\chi^2$	$df$	$p$	$V$
	$n (%)$	$n (%)$	$n (%)$				
Gender				.59	2	.745	.04
Female	50 (51.0)	54 (51.4)	48 (46.6)				
Male	48 (49.0)	51 (48.6)	55 (53.4)				

Note. Column percentages.  $n$  = absolute frequency, % = relative frequency,  $\chi^2$  = Chi-square,  $df$  = degrees of freedom,  $p$  = significance level,  $V$  = Cramer's  $V$ . "Other" excluded from the analysis.

**Table A9**  
Randomization Check: Education (Aggregation Criteria)

	Arguments $n = 100$	Opinions $n = 105$	Facts $n = 106$	$\chi^2$	$df$	$p$	$V$
	$n$ (%)	$n$ (%)	$n$ (%)				
Education				11.91	6	.064	.06
No degree	0 (.0)	0 (.0)	0 (.0)				
Lower secondary school degree	9 (9.0)	5 (4.8)	11 (10.4)				
Intermediate secondary school degree	27 (27.0)	47 (44.8)	32 (30.2)				
(Technical) university entrance qualification	22 (22.0)	26 (24.8)	25 (23.6)				
University degree	42 (42.0)	27 (25.7)	38 (35.8)				

Note. Column percentages.  $n$  = absolute frequency, % = relative frequency,  $\chi^2$  = Chi-square,  $df$  = degrees of freedom,  $p$  = significance level,  $V$  = Cramer's  $V$ . "Still in school" excluded from the analysis.

## Data availability

Data will be made available on request.

## References

- Abdulrauf, A. A., Hamid, N. A., & Ishak, M. S. (2017). Cognitive engagement and online political participation on social media among youths in Malaysia: The moderating role of political knowledge. In *SHS web of conferences* (Vol. 33), Article 00085. <https://doi.org/10.1051/shsconf/20173300085>
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13(3), 219–235. <https://doi.org/10.1177/1088868309341564>
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "nasty effect": Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3), 373–387. <https://doi.org/10.1111/jcc4.12009>
- Backstrom, L., Kleinberg, J., Lee, L., & Danescu-Niculescu-Mizil, C. (2013). Characterizing and curating conversation threads: Expansion, focus, volume, re-entry. In *Proceedings of the sixth ACM international conference on web search and data mining* (pp. 13–22). Association for Computing Machinery. <https://doi.org/10.1145/2433396.2433401>
- Banducci, S., Giebler, H., & Kritzing, S. (2017). Knowing more from less: How the information environment increases knowledge of party positions. *British Journal of Political Science*, 47(3), 571–588. <https://doi.org/10.1017/S0007123415000204>
- Bhaskar, A., Fabbri, A., & Durrett, G. (2023). Prompted opinion summarization with GPT-3.5. In *Findings of the association for computational linguistics: Acl 2023* (pp. 9282–9300). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.findings-acl.591>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64(1), 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Black, L. W., Welsler, H. T., Cosley, D., & DeGroot, J. M. (2011). Self-governance through group discussion in wikipedia: Measuring deliberation in online groups. *Small Group Research*, 42(5), 595–634. <https://doi.org/10.1177/1046496411406137>
- Campbell, M., Julious, S., & Altman, D. (1995). Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two-group comparisons. *BMJ*, 311(7013), 1145–1148. <https://doi.org/10.1136/bmj.311.7013.1145>
- Cao, M., Dong, Y., & Cheung, J. C. K. (2022). Hallucinated but factual! inspecting the factuality of hallucinations in abstractive summarization. In *Proceedings of the 60th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 3340–3354). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.236>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Chen, Y.-C., Shang, R.-A., & Kao, C.-Y. (2009). The effects of information overload on consumers' subjective state toward buying decision in the internet shopping environment. *Electronic Commerce Research and Applications*, 8(1), 48–58. <https://doi.org/10.1016/j.elerap.2008.09.001>
- Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (pp. 1217–1230). Association for Computing Machinery. <https://doi.org/10.1145/2998181.2998213>
- Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, 64(4), 658–679. <https://doi.org/10.1111/jcom.12104>
- Dalisy, F., Hmielowski, J. D., Kushin, M. J., & Yamamoto, M. (2012). Social capital and the spiral of silence. *International Journal of Public Opinion Research*, 24(3), 325–345. <https://doi.org/10.1093/ijpor/eds023>
- Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press.
- Diakopoulos, N., & Naaman, M. (2011). Towards quality discourse in online news comments. In *Proceedings of the ACM 2011 conference on computer supported cooperative work* (pp. 133–142). Association for Computing Machinery. <https://doi.org/10.1145/1958824.1958844>
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83–87. <https://doi.org/10.1111/1467-8721.01235>
- Edwards, P., Hindmarsh, R., Mercer, H., Bond, M., & Rowland, A. (2008). A three-stage evaluation of a deliberative event on climate change and transforming energy. *Journal of Public Deliberation*, 4(1). <https://doi.org/10.16997/jdd.65>. Article 1.
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society*, 20(5), 325–344. <https://doi.org/10.1080/01972240490507974>
- Esau, K., Wilms, L., Bales, J., & Keller, B. (2023). For deliberation sake, show some constructive emotion! how different types of emotions affect the deliberative quality of interactive user comments. *Javnost – The Public*, 30(4), 472–495. <https://doi.org/10.1080/13183222.2023.2171217>
- Eveland, W. P., & Hutchens, M. J. (2011). Political knowledge. In *The international encyclopedia of communication* (2nd ed.). Wiley. <https://doi.org/10.1002/9781405186407.wbiecp060.pub2>
- Faridani, S., Bitton, E., Ryokai, K., & Goldberg, K. (2010). Opinion space: A scalable tool for browsing online user reviews. In *Proceedings of the 28th international conference on human factors in computing systems* (pp. 1175–1184). Association for Computing Machinery. <https://doi.org/10.1145/1753326.1753502>
- Flynn, L. R., & Goldsmith, R. E. (1999). A short, reliable measure of subjective knowledge. *Journal of Business Research*, 46(1), 57–66. [https://doi.org/10.1016/S0148-2963\(98\)00057-5](https://doi.org/10.1016/S0148-2963(98)00057-5)
- Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, 8(1), 10–16.
- Friess, D., & Eilders, C. (2015). A systematic review of online deliberation research. *Policy & Internet*, 7(3), 319–339. <https://doi.org/10.1002/poi3.95>
- Friess, D., Ziegele, M., & Heinbach, D. (2021). Collective civic moderation for deliberation? Exploring the links between citizens' organized engagement in comment sections and the deliberative quality of online discussions. *Political Communication*, 37(5), 643–665. <https://doi.org/10.1080/10584609.2020.1830322>
- Gao, M., Do, H. J., & Fu, W.-T. (2017). An intelligent interface for organizing online opinions on controversial topics. In *Proceedings of the 22nd international conference on intelligent user interfaces* (pp. 111–120). Association for Computing Machinery. <https://doi.org/10.1145/3025171.3025230>
- Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, 10(6), 597–602. <https://doi.org/10.3758/BF03202442>
- Graf, B., & Antoni, C. H. (2020). The relationship between information characteristics and information overload at the workplace: A meta-analysis. *European Journal of Work & Organizational Psychology*, 30(1), 1–16. <https://doi.org/10.1080/1359432X.2020.1813111>
- Gudowsky, N., & Bechthold, U. (2013). The role of information in public participation. *Journal of Public Deliberation*, 9(1). <https://doi.org/10.16997/jdd.152>. Article 3.
- Habermas, J. (1992). Further reflections on the public sphere. In C. Calhoun (Ed.), *Habermas and the public sphere* (pp. 421–461). MIT Press.
- Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy*. MIT Press.
- Hargittai, E., Neuman, W. R., & Curry, O. (2012). Taming the information tide: Perceptions of information overload in the American home. *The Information Society*, 28(3), 161–173. <https://doi.org/10.1080/01972243.2012.669450>
- Hayes, A. F. (2021). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). Guilford Press.
- Heath, R. L., & Douglas, W. B. (1990). Involvement: A key variable in people's reaction to public policy issues. *Public Relations Research Annual*, 2(1–4), 193–204. [https://doi.org/10.1207/s1532754xjpr0201-4\\_8](https://doi.org/10.1207/s1532754xjpr0201-4_8)
- Hsu, C.-F., Khabiri, E., & Caverlee, J. (2009). Ranking comments on the social web. In *Proceedings of the international conference on computational science and engineering* (pp. 90–97). IEEE. <https://doi.org/10.1109/CSE.2009.109>

- Hu, H., & Krishen, A. S. (2019). When is enough, enough? Investigating product reviews and information overload from a consumer empowerment perspective. *Journal of Business Research*, 100, 27–37. <https://doi.org/10.1016/j.jbusres.2019.03.011>
- Hubler, J. N. (2021). Aristotle and the complexities of opinion. In J. N. Hubler (Ed.), *Overcoming uncertainty in ancient Greek political philosophy* (pp. 81–98). Palgrave Macmillan. [https://doi.org/10.1007/978-3-030-82091-6\\_4](https://doi.org/10.1007/978-3-030-82091-6_4).
- In der Au, A.-M., Wieland, M., & Zeitler, A.-L. (2017). Zufällig gut informiert? Folgen gezielter und zufälliger Nachrichtenkontakte für politisches Wissen und subjektive Informiertheit [Randomly well-informed? Effects of targeted and incidental news exposure on political knowledge and perceived knowledgeability]. *MedienJournal*, 41(2), 61–75. <https://doi.org/10.24989/medienjournal.v41i2.1471>
- Janssen, D., & Kies, R. (2005). Online forums and deliberative democracy. *Acta Politica*, 40(3), 317–335. <https://doi.org/10.1057/palgrave.ap.5500115>
- Jhaver, S., Appling, D. S., Gilbert, E., & Bruckman, A. (2019). “did you suspect the post would be removed?”: Understanding user reactions to content removals on Reddit. In *Proceedings of the ACM on human-computer interaction* (Vol. 3) Association for Computing Machinery. <https://doi.org/10.1145/3359294> (CSCW), Article 192.
- Jones, Q., & Rafaeli, S. (1999). User population and user contributions to virtual publics. In *Proceedings of the international ACM SIGGROUP conference on supporting group work* (pp. 239–248). Association for Computing Machinery. <https://doi.org/10.1145/320297.320325>.
- Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research*, 15(2), 194–210. <https://doi.org/10.1287/isre.1040.0023>
- Kim, J. W., Guess, A., Nyhan, B., & Reifler, J. (2021). The distorting prism of social media: How self-selection and exposure to incivility fuel online comment toxicity. *Journal of Communication*, 71(6), 922–946. <https://doi.org/10.1093/joc/jqab034>
- Kim, Y., & Khang, H. (2014). Revisiting civic voluntarism predictors of college students' political participation in the context of social media. *Computers in Human Behavior*, 36, 114–121. <https://doi.org/10.1016/j.chb.2014.03.044>
- Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, 22(6), 787–794. <https://doi.org/10.1177/0956797611407929>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Ksiazek, T. B. (2018). Commenting on the news: Explaining the degree and quality of user comments on news websites. *Journalism Studies*, 19(5), 650–673. <https://doi.org/10.1080/1461670X.2016.1209977>
- Ksiazek, T. B., & Springer, N. (2018). User comments in digital journalism: Current research and future directions. In S. Eldridge II, & B. Franklin (Eds.), *The routledge handbook of developments in digital journalism studies* (pp. 475–486). Routledge.
- Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). *Strategies of knowledge acquisition*. Society for Research in Child Development. <https://doi.org/10.2307/1166059>
- Kwon, B. C., Kim, S.-H., Duket, T., Catalán, A., & Yi, J. S. (2015). Do people really experience information overload while reading online reviews? *International Journal of Human-Computer Interaction*, 31(12), 959–973. <https://doi.org/10.1080/10447318.2015.1072785>
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Lee, E.-J., & Jang, Y. J. (2010). What do others' reactions to news on internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research*, 37(6), 825–846. <https://doi.org/10.1177/0093650210376189>
- Legrand, M., Kessler, S. H., & Eisenreich, C. (2019). Politische Inhalte liken, teilen und im Internet lesen – Zum Begriffsverständnis von politischer Online-Partizipation [Liking, sharing, and reading political content online: On the conceptual understanding of political online participation]. In I. Engelmann, M. Legrand, & H. Marzinkowski (Eds.), *Politische Partizipation im Medienwandel* (pp. 53–76). <https://doi.org/10.17174/dcr.v6.3>. Nomos.
- Li, H., Schlegel, V., Sun, Y., Batista-Navarro, R., & Nenadić, G. (2025). *Large language models in argument mining: A survey* [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2506.16383>
- Lim, H. (2008). The use of different happiness rating scales: Bias and comparison problem. *Social Indicators Research*, 87, 259–267. <https://doi.org/10.1007/s11205-007-9171-x>
- Llewellyn, C., Grover, C., & Oberlander, J. (2016). Improving topic model clustering of newspaper comments for summarisation. In *Proceedings of the 54th annual meeting of the association for computational linguistics: Student research workshop* (pp. 43–50). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P16-3007>
- Ma, Z., Sun, A., Yuan, Q., & Cong, G. (2012). Topic-driven reader comments summarization. In *Proceedings of the 21st ACM international conference on information and knowledge management* (pp. 265–274). Association for Computing Machinery. <https://doi.org/10.1145/2396761.2396798>
- Matias, J. N. (2019). Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 116(20), 9785–9789. <https://doi.org/10.1073/pnas.1813486116>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511617201>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Nematzadeh, A., Ciampaglia, G., Ahn, Y.-Y., & Flammini, A. (2019). Information overload in group communication: From conversation to cacophony in the Twitch chat. *Royal Society Open Science*, 6. <https://doi.org/10.1098/rsos.191412>. Article 191412.
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283. <https://doi.org/10.1177/146144804041444>
- Park, D.-H., Lee, J., & Han, I. (2006). Information overload and its consequences in the context of online consumer reviews. In *In proceedings of the Pacific Asia conference on information systems (PACIS 2006)*. Association for Information Systems.
- Peacock, C., Scacco, J. M., & Stroud, N. J. (2019). The deliberative influence of comment section structure. *Journalism*, 20(6), 752–771. <https://doi.org/10.1177/1464884917711791>
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer.
- Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of civility and reasoning in user comments on perceived journalistic quality. *Journalism Studies*, 19(1), 62–78. <https://doi.org/10.1080/1461670X.2016.1161497>
- Ran, W., Yamamoto, M., & Xu, S. (2016). Media multitasking during political news consumption: A relationship with factual and subjective political knowledge. *Computers in Human Behavior*, 56, 352–359. <https://doi.org/10.1016/j.chb.2015.12.015>
- Rios, K., Goldberg, M. H., & Totton, R. R. (2017). An informational influence perspective on (non)conformity: Perceived knowledgeability increases expression of minority opinions. *Communication Research*, 45(2), 241–260. <https://doi.org/10.1177/0093650217699935>
- Roush, A., & Balaji, A. (2020). DebateSum: A large-scale argument mining and summarization dataset. In *Proceedings of the 7th workshop on argument mining* (pp. 1–7). Association for Computational Linguistics. <https://aclanthology.org/2020.argmining-1.1/>
- Roush, A., Shabazz, Y., Balaji, A., Zhang, P., Mezza, S., Zhang, M., Basu, S., Vishwanath, S., Fatemi, M., & Shwartz-Ziv, R. (2024). OpenDebateEvidence: A massive-scale argument mining and summarization dataset. *arXiv*. <https://doi.org/10.48550/arXiv.2406.14657> [Preprint].
- Ruiz, C., Domingo, D., Micó, J. L., Diaz-Noci, J., Masip, P., & Meso, K. (2011). Public sphere 2.0? The democratic qualities of citizen debates in online newspapers. *The International Journal of Press/Politics*, 16(4), 463–487. <https://doi.org/10.1177/1940161211415849>
- Ryffel, F. A., & Wirth, W. (2020). How perceived processing fluency influences the illusion of knowing in learning from TV reports. *Journal of Media Psychology*, 32(1), 2–13. <https://doi.org/10.1027/1864-1105/a000250>
- Salmon, C. T., & Neuwirth, K. (1990). Perceptions of opinion climates and willingness to discuss the issue of abortion. *Journalism Quarterly*, 67(3), 567–577. <https://doi.org/10.1177/107769909006700312>
- Schaefer, S. (2022a). *Mediennutzung und metakognitive Urteile: Die Rolle digitaler Nachrichten für die Wahrnehmung von Wissen [Media use and metacognitive judgments: The role of digital news in the perception of knowledge]*. (Digital Communication Research, 11). Weizenbaum Institute. <https://doi.org/10.48541/dcr.v11.0>
- Schaefer, S., Müller, P., & Ziegele, M. (2022b). The double-edged sword of online deliberation: How evidence-based user comments both decrease and increase discussion participation intentions on social media. *New Media & Society*, 26(3), 1403–1428. <https://doi.org/10.1177/1461448211073059>
- Schneider, S. M. (1997). *Expanding the public sphere through computer-mediated communication: Political discussion about abortion in a Usenet newsgroup (doctoral dissertation)*. Massachusetts Institute of Technology.
- Schneider, F. M., Weinmann, C., Roth, F. S., Knop, K., & Vorderer, P. (2016). Learning from entertaining online video clips? Enjoyment and appreciation and their differential relationships with knowledge and behavioral intentions. *Computers in Human Behavior*, 54, 475–482. <https://doi.org/10.1016/j.chb.2015.08.028>
- Seering, J., Kraut, R., & Dabbish, L. (2017). Shaping pro and anti-social behavior on Twitch through moderation and example-setting. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (pp. 111–125). Association for Computing Machinery. <https://doi.org/10.1145/2998181.2998277>
- Shamir, J. (1997). Speaking up and silencing out in face of a changing climate of opinion. *Journalism & Mass Communication Quarterly*, 74(3), 602–614. <https://doi.org/10.1177/107769909707400313>
- Springer, N., Engelmann, I., & Pfaffinger, C. (2015). User comments: Motives and inhibitors to write and read. *Information, Communication & Society*, 18(7), 798–815. <https://doi.org/10.1080/1369118X.2014.997268>
- Strandberg, K., & Berg, J. (2013). Online newspapers' readers' comments – Democratic conversation platforms or virtual soapboxes? *Comunicação e Sociedade*, 23(1), 132–152. [https://doi.org/10.17231/comsoc.23\(2013\).1618](https://doi.org/10.17231/comsoc.23(2013).1618)
- Stromer-Galley, J. (2007). Measuring deliberation's content: A coding scheme. *Journal of Public Deliberation*, 3(1). <https://doi.org/10.16997/jdd.50>. Article 12.
- Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication*, 20(2), 188–203. <https://doi.org/10.1111/jcc4.12104>
- Stroud, N. J., Van Duyn, E., & Peacock, C. (2016). *News commenters and news comment readers* [report]. Engaging news project. <https://mediaengagement.org/wp-content/uploads/2016/03/ENP-News-Commenters-and-Comment-Readers1.pdf>.

- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185–233. [https://doi.org/10.1207/s1532690xcil203\\_1](https://doi.org/10.1207/s1532690xcil203_1)
- Tenenboim, O., & Cohen, A. (2015). What prompts users to click and comment: A longitudinal study of online news. *Journalism*, 16(2), 198–217. <https://doi.org/10.1177/1464884913513996>
- Towne, W. B., & Herbsleb, J. D. (2012). Design considerations of online deliberation systems. *Journal of Information Technology & Politics*, 9(1), 97–115. <https://doi.org/10.1080/19331681.2011.637711>
- Uttl, B. (2005). Measurement of individual differences. *Psychological Science*, 16(6), 460–467. <https://doi.org/10.1111/j.0956-7976.2005.01557.x>
- Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research*, 43(3), 476–496. <https://doi.org/10.1080/00273170802285941>
- Weber, P. (2014). Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media & Society*, 16(6), 941–957. <https://doi.org/10.1177/1461444813495165>
- Winter, S., Brückner, C., & Krämer, N. C. (2015). They came, they liked, they commented: Social influence on Facebook news channels. *Cyberpsychology, Behavior, and Social Networking*, 18(8), 431–436. <https://doi.org/10.1089/cyber.2015.0005>
- Wright, S., & Street, J. (2007). Democracy, deliberation and design: The case of online discussion forums. *New Media & Society*, 9(5), 849–869. <https://doi.org/10.1177/1461444807081230>
- Yun, G. W., & Park, S.-Y. (2011). Selective posting: Willingness to post a message online. *Journal of Computer-Mediated Communication*, 16(2), 201–227. <https://doi.org/10.1111/j.1083-6101.2010.01533.x>
- Ziegele, M., Quiring, O., Esau, K., & Friess, D. (2020). Linking News Value Theory With Online Deliberation: How News Factors and Illustration Factors in News Articles Affect the Deliberative Quality of User Discussions in SNS' Comment Sections. *Communication Research*, 47(6), 860–890. <https://doi.org/10.1177/0093650218797884>